

TÜRKÇE METİNLERİN DİNAMİK ANALİZİ

Gökhan Şahin, Murat Erentürk ve Avadis Hacınlıyan

Yeditepe Üniversitesi, Bilişim Sistemleri ve Fizik Bölümleri,
İnönü Mah. Kayışdağı Cad.
26 Ağustos Yerleşimi 34755 Kadıköy, İstanbul
sahin@yeditepe.edu.tr
merenturk@yeditepe.edu.tr
ahacinliyan@yeditepe.edu.tr

ÖZET

Bu çalışmada kelime sıklığı dağarlarına dayanmayan yeni bir bağımlı değişken önerilmiş, bu parametrizasyon yardımı ile değişik doğal dillerdeki metinlerden zaman serisi oluşturulmuş, uzun erimli korelasyon ve fraktal yapıların var olabileceği gösterilmiştir. Üssel korelasyonlarını betimleyen bir ölçeklenme katsayısı bulmak için eğilimlerden arındırılmış dalgalanma analizi ve Difüzyon Entropisi Analizi kullanılmış, bu iki yöntemin iki değişik rejimi içeren uyumlu sonuç verdiği ve değişik dillerin bu analizle betimlenebileceği ve birbirinden ayırt edilebileceği gösterilmiştir.

ABSTRACT

In this work, a new dependent variable that is not based on word frequency corpora has been proposed. Using this parameterization, time series have been formed from texts in natural languages, the existence of long range correlations and fractal structures has been demonstrated. A scaling parameter that determines the long range exponential correlations has been sought both by Detrended Fluctuation Analysis-Dfa and Diffusion Entropy analysis. Results from both analyses are in agreement and reveal two distinct regimes that can characterize and distinguish different languages.

1. GİRİŞ

Doğadaki birçok yapının matematikte fraktaller adı verilen tanıma uyduğu gözlemlenmiştir. Fraktaller uzayda kendine benzer yapılardan oluşan, zaman evrimi sırasında pek çok sistem için $1/f$ gürültüsü olarak bilinen davranışı gösteren, kesirli boyutlu ve evrensellik gösteren yapılardır. Günümüzde, yeni gelişmekte olan bu kavramın, doğal dillerin iç yapısını ve eğer mevcut ise çeşitli dil gurupları arasındaki farklılık derecelerini anlayabilmek üzere kullanılabilecek bir araç olduğu ileri sürülmektedir. Verilen bir dildeki yazılı metinler üzerinde hece, kelime, cümle, vb. gibi değişik katmanlarda yapılacak zaman serisi veya dalgalanma analizlerinin birbirine yakın sonuçlar vererek bu dilin değişik ölçeklerde kendi kendine benzeyen bir karakter sergileyeceği manasına gelir. Ayrıca, değişik diller de birbirinden ayırt edilebilir. Doğal diller bir düzeyler hiyerarşisi şeklindedir. En alt düzey olan hece ve harfler seslendirme bakımından önemli, ancak anlam bakımından ilginç değildir. Anlam kelime ve bunun üstündeki düzeylerde ortaya çıkar ve bu düzey dilbilim açısından önemlidir. Bu değişik yapıların sınıflandırılması istatistiksel yöntemlerin kontrollü ve sistematik şekilde dilbilim verilerine uygulanmasını gerektirir. Dilde rassallık bakımından yapılan ilk çalışmalar değişik dilbilimsel formların istatistiksel sıklıklarını gösteren frekans tablolarının oluşturulmasıdır. Evrim sürecine uyan değişik yapılar gibi doğal diller de yazılı bir metin olarak uzayda, konuşma sırasında seslerin arka arkaya gelmesi bakımından kısa süreli olarak zamanda, evrim bakımından da çok uzun süreli olarak zamanda karmaşık ve çok ölçekli bir yapıya sahiptirler. Bilgisayar bilimlerindeki ve teknolojiye hızlı ilerlemeler in sunduğu araştırma olanağı nedeniyle Türkçe gibi doğal diller enformatik, şifreleme, telekomünikasyon ve matematik gibi bilim dallarının ilgi alanı haline gelmiştir. İlk aşamada başvuru istatistiksel yöntemler sistemin ortalama davranışını ve bu davranış etrafındaki sapmaları incelediğinden sistemin uzun vadeli davranışı konusunda bu yaklaşım ile yeterince bilgi edinemeyiz. Bu nedenle, lineer istatistik yöntemlerine ek olarak değişik araştırmacılar, doğal dillerde kendine benzer fraktal yapıların yer aldığını da ileri sürmüşlerdir. Örnek olarak Montemurro ve Purry [1] Shakespeare'in 36 oyununda bu tür bir davranış gözlemlemişlerdir. Bahm ve başkaları bu durumun İngilizce dilinin mi Shakespeare'in üslubuna ait bir özellik olup olmadığını araştırmak için aynı analizi hem bu oyunların Kore diline tercümelemleri hem de Koreli yazarların popüler romanları üzerinde gerçekleştirmişlerdir [2]. Öte yandan üslup determinizm ilkesinin her zaman uygulanabileceği bir özellik değildir. Örneğin kullanılan kelime sayısında İncil'de 5649, Dante'de 5860 (bunlardan 1615'i özel isim), Milton'da 8000 kadar, Shakespeare'de 15000 şeklinde önemli farklılıklar vardır.

Daha dikkatli bir analiz, ortalama davranıştan uzun vadede sistematik bir hal alan sapmalar yeniden ölçümlendirilen erim (Rescaled range) gibi lineer olmayan analiz yöntemleri ile gerçekleştirilebilir. İlk kez Hurst tarafından Nil nehrinin taşma verileri için gerçekleştirilen bu tür bir inceleme karmaşıklık ölçütlerini belirlemeye yarar [3]. Doğal dillerde hem rastsal yapılar, hem de bir sonraki adımın bir önceki ile bağımlı olması şeklindeki rasgele yürüyüşü anımsatan bir düzen vardır. Bu amaçla yapılan ilk çalışmalar değişik dilbilimsel formların istatistiksel sıklıklarını gösteren sıklık tablolarının (corpus) oluşturulması ve üzerinde istatistiksel analizlerin yapılmasıdır. Değişik metinlerden alınan örneklerde bile belirli bir istikrar gözlenmiştir.

Kısa vadeli sapmalar ise bilgi kuramına dayalı ölçütlerle incelenebilir. Bu ölçütler bir yandan verinin rastsallığını, diğer yandan ise karmaşıklığını dikkate alır. Karmaşıklık sabit, tümüyle periyodik veya tümüyle rastsal veriler için düşük, bu basit davranışlarla açıklanamayan veri yapıları için yüksek değerler alan bir parametre ile niteleyebiliriz

Böylesi bir inceleme bağımsız bir değişkene ihtiyaç gösterir. Batı dilleri üzerinde bu yönde yapılan birçok çalışma vardır. İlk eserler kelime sıklıklarının istatistiksel analizine dayanır ve Zipf tarafından öne sürülen dört kuralın irdelenmesini konu alır. Eğer bir dil bir dizi kelime ile bunların kullanım sıklıkları olarak yorumlanırsa bir istatistiksel dağılım için gerekli öğeler sağlanmış olur. Eğer dil, düşünceleri kodlayan bir sistemse buna da “bilgi kuramı”na ait yaklaşımları uygulayabiliriz. Bu amaçla kelime veya hece uzunluklarının Hurst analizi ile fraktal yapıların incelenmesine yönelik çalışmalar gerçekleştirilmiştir. Buna bir örnek olarak Hrebicek [4,5,6] tarafından gerçekleştirilen ve Türkçe’yi de içeren çalışma verilebilir.

Türkçe, eklerden geniş surette yararlanması ve kurallı bir dilbilgisine sahip olması nedenleriyle fraktal analize yanıt veren bir dinamik yapı oluşturmaya adaydır. Ancak bu durumun yarattığı bir güçlük de vardır: aynı kökten olağanüstü çok sayıda sözcük türetilir.

Uzun zamanlı korelasyonların varlığını bulmak için sistemin zamandaki evrimini veren bağımlı bir değişkene ihtiyaç vardır. Metinler üzerinde şimdiye kadar yapılan tüm çalışmalarda bu amaçla dağarlardan (corpus) yararlanılmıştır [1,2, 7]. Bundan dolayı çalışmalar en fazla aynı anda iki dile sınırlandırılmış ve dağara bağımlı olmuşlardır. Bizim bu bağlamda literatürde bulabildiğimiz 3 dille ilgili çalışma vardır İtalyanca, Korece, İngilizce. Tüm bu çalışmalar diğer dillerde uzun vadeli korelasyonların varlığını işaret etmekle birlikte diğer diller için kesin bir sonuç gösterememektedirler.

Bu noktada dağardan bağımsız bir araştırma yöntemine ihtiyaç doğmaktadır. Bunun için DNA Rastsal Yürüyüşlerinden esinlenen yeni bir yöntem önermekteyiz. [8]. Yöntemde her sözcük rastsal yürüyüşte bir adım olarak ele alınır ve uzunluğu

$$L(N) = \sum_{i=1}^N y(i)$$
 bağıntısı ile verilmektedir. Burada N kelimenin uzunluğunu y(i) ise i. sıradaki harfin unicode değerini verir. İnceleme konusu olan metinde ki her sözcük rastsal yürüyüşte ki bir adıma çevrilir. Sözcüklerin metinde ki sıralaması ise o adımın zamanı olarak ele alınır.

Metinlerden zaman serisi oluşturulduktan sonraki aşama, zaman serilerinde ki uzun zamanlı üssel korelasyonlarını betimleyecek bir ölçeklenme katsayısı bulmaktır. Bu çalışmada bunun için iki yöntem kullanılmıştır eğilimlerden arındırılmış dalgalanma analizi (Detrended Fluctuation Analysis-Dfa) [9] ve Difüzyon Entropisi Analizi [10].

Eğilimlerden arındırılmış dalgalanma analizi durağan olmayan zaman serilerindeki uzun zamanlı üssel korelasyon katsayılarını bulmak için kullanılan bir ölçeklenme analiz yöntemidir. 1994 yılında ilk ortaya atıldığından bu yana DNA nükleotidlerinin organizasyonundan kalp atışı zaman serilerine, ekonomi zaman serilerinden, hava tahmini, katı hal fizikine kadar bir çok alanda yer bulmuştur [11-20].

Difüzyon Entropisi Analizi ise görece daha yeni bir yöntemdir. Diğer yöntemlerin aksine her zaman doğru ölçeklenme katsayısını verdiği iddia edilmektedir[21,22].

2. EĞİLİMDEN ARINDIRILMIŞ DALGALANMA ANALİZİ

Durağanlaşmamış (nonstationary) zaman serilerinde ki korelasyon özelliklerini nitelendirilmek için kullanılan bir yöntemdir. Zaman serisindeki ölçeklendirme üstelinin (bundan sonra α ile belirtilecek) modifiye edilmiş karekök ortalama (root mean square) yöntemiyle hesaplanmasına dayanır [2].

α 'yı $x(i)$ [$i=1, \dots, N$] ile gösterilen bir zaman serisinden hesaplamak için önce zaman serisi aşağıda ki gibi entegre edilir:

$$y(k) = \sum_{i=1}^k x(i) - \bar{x}.$$

Burada \bar{x} zaman serisinin ortalama değeridir. k ise 1 ile N arasında değişmektedir.

Daha sonra $y(k)$ n uzunluğunda eşit kutulara bölünür. Her kutudaki veriye en küçük kare (least squares) yöntemiyle bir doğru uydurulur ($y_n(k)$). Daha sonra entegre edilmiş zaman serisi yerel eğim ($y_n(k)$) çıkarılmak suretiyle eğilimlerinden arındırılır. Eğilimden arındırılmış zaman serilerinin karekök ortalama (root mean square fluctuation) dalgalanması, ($F(n)$) şöyle hesaplanır:

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n(k)]^2}$$

$F(n)$ tüm n değerleri için hesaplanır. $\log(F(n))$, $\log(n)$ grafiğinin eğimi α 'yı verir. α ile $1/f$ eğimi arasında $1/f$ eğimi $= 2\alpha - 1$ şeklinde bir ilişki vardır:

3. DİFÜZYON ENTROPİSİ ANALİZİ

M sayıdan oluşan ξ_i , $i=1, \dots, M$. dizisini varsayalım, Difüzyon Entropisi Analizinin amacı normal veya anormal dizide var olan ölçeklenmeyi verilere hiç bir eğilimden arındırma işlemi uygulamadan bulmaktır. İlk olarak $1 \leq l \leq M$ koşuluna uyan bir l sayısı seçilir. Bu sayı bundan sonra ‘zaman’ olarak adlandırılacaktır. Verilen her hangi bir l zamanı için $\xi_i^{(s)} \equiv \xi_{i+s}$, $s=0, \dots, M-l$ koşuluna uyan $M-l+1$ tane alt dizi bulunabilir. Bu alt dizilerden her hangi biri için konumu s ile tanımlanan

$$x^{(s)}(l) = \sum_{i=1}^l \xi_i^{(s)} = \sum_{i=1}^l \xi_{i+s}.$$

bir difüzyon yörüngesi (trajectory) inşa edilebilir. Bu pozisyon düzenli aralıklarla yukarıda ki tarife göre ilgili alt dizi için ileri veya geri zıplayan bir Brown parçacığının pozisyonu olarak düşünülebilir. Bu parçacığın l zamanında bulunduğu yere gelmek için l defa zıplama yaptığı anlamına gelir. Her hangi bir i adımı yapılan zıplamanın yoğunluğu

$|\xi_i^{(s)}|$ olacaktır, zıplamanın ileriye mi yoksa geriye mi olduğu ise $\xi_i^{(s)}$ 'in işareti tarafından belirlenecektir.

Artık bu difüzyon işleminin entropisi hesaplanabilir. Bunun için x eksenini l büyüklüğünde hücrelere bölünür. Her hücre isimlendirildikten sonra her hangi bir l anında bir hücrede bulunan parçacıklar sayılır. Bu sayıyı $N_i(l)$ ile gösterelim. Böylece her hangi bir parçacığın l zamanında i . hücrede olma olasılığı

$$p_i(l) \equiv \frac{N_i(l)}{(M-l+1)}$$

ile, bu durumda her hangi bir l anında ki difüzyon işleminin entropisi

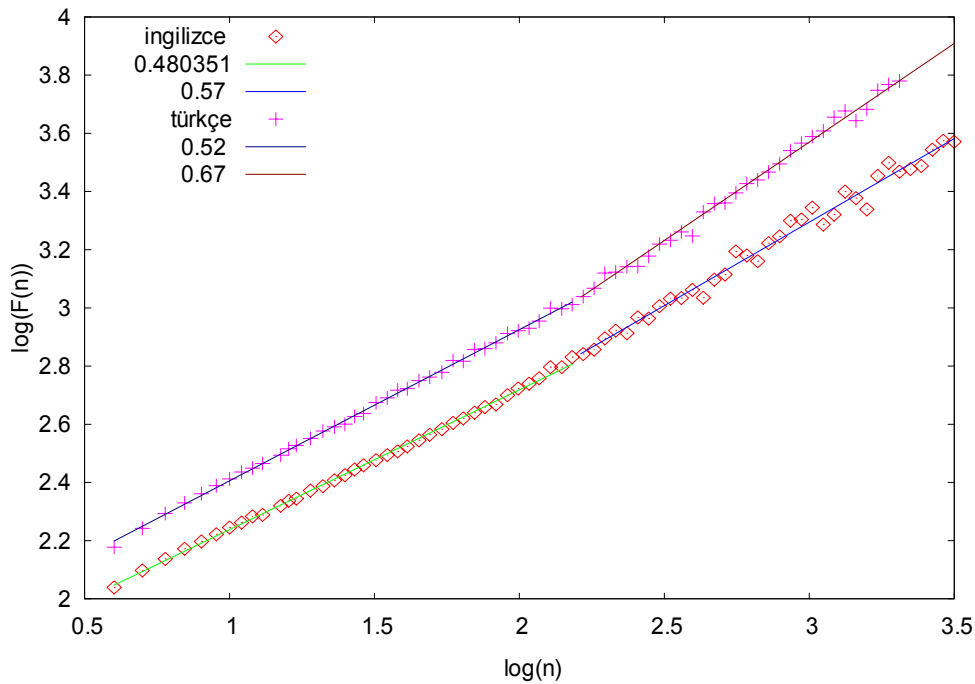
$$S_d(l) = -\sum_i p_i(l) \ln[p_i(l)]$$

ile verilir. Tüm zaman serisi için elde edilecek olan $S_d(l)$ - l grafiğinin eğimi (δ) bize ölçeklenme üsselini verecektir. Eğer sürecin temelini oluşturan difüzyon Gaussyan ise Eğimden Arındırılmış Dalgalanma Analizinin vereceği üssel α ile δ arasında $\delta = \alpha$ ilişkisinin olduğu, yok eğer Levy istatistiği ile belirtiliyorsa bu durumda üsseller arasında

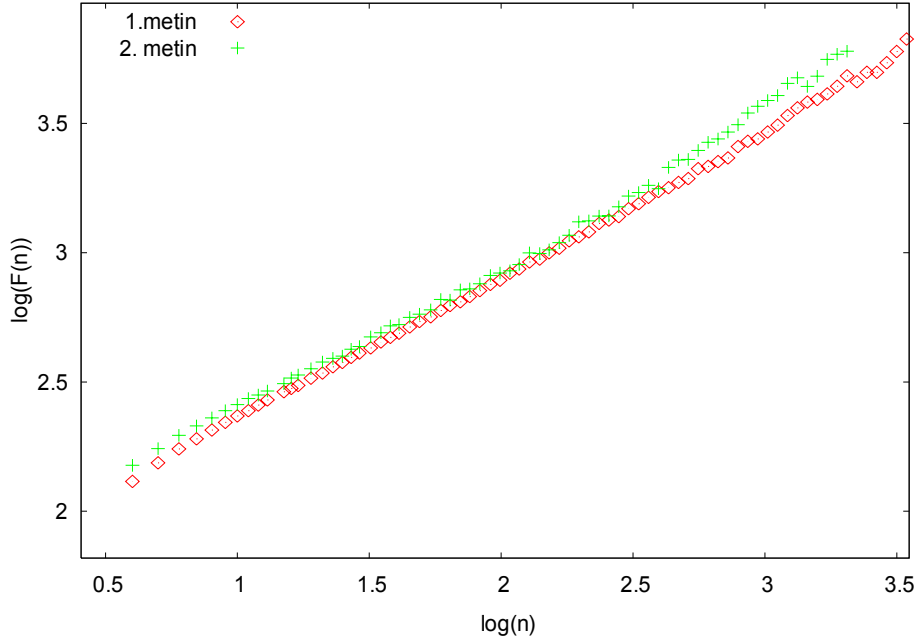
$$\delta = \frac{1}{3-2\alpha}$$
 ilişkisinin olduğu gösterilmiştir [23].

4. ANALİZ

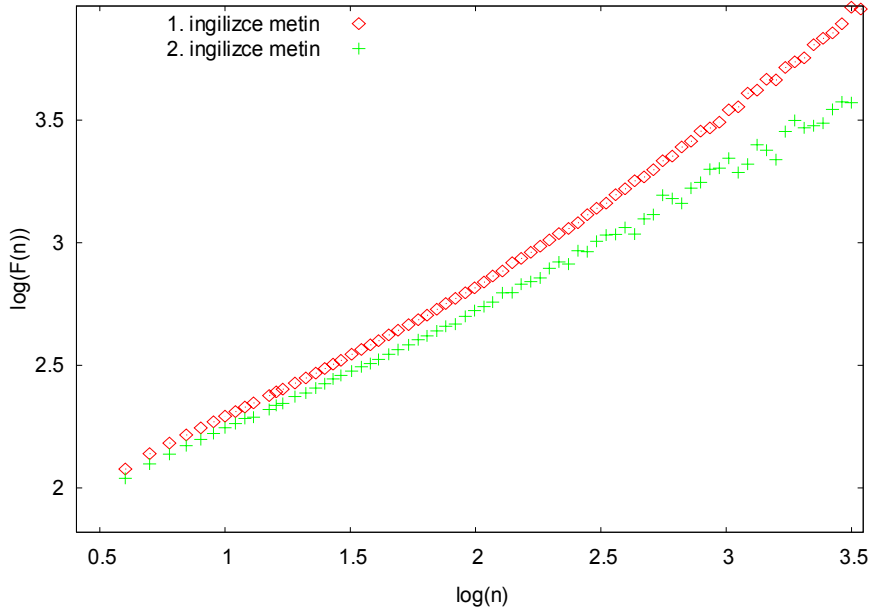
Analiz için ikisi aynı metnin İngilizce ve Türkçesi olmak üzere dört tane metin kullanılmıştır. Aşağıda ki grafikte bir birinin tercümesi olan iki (biri Türkçe diğeri İngilizce) metnin eğilimden arındırılmış dalgalanma analizi (DFA) görülmektedir. Eğimlerdeki kırılmalar dillerin korrelasyon özelliklerindeki değişimi göstermektedir. Eğimlerin değerleri grafiğin üstünde gösterilmiştir. İki dil arasındaki korelasyon farklılıkları gözlenebilmektedir.



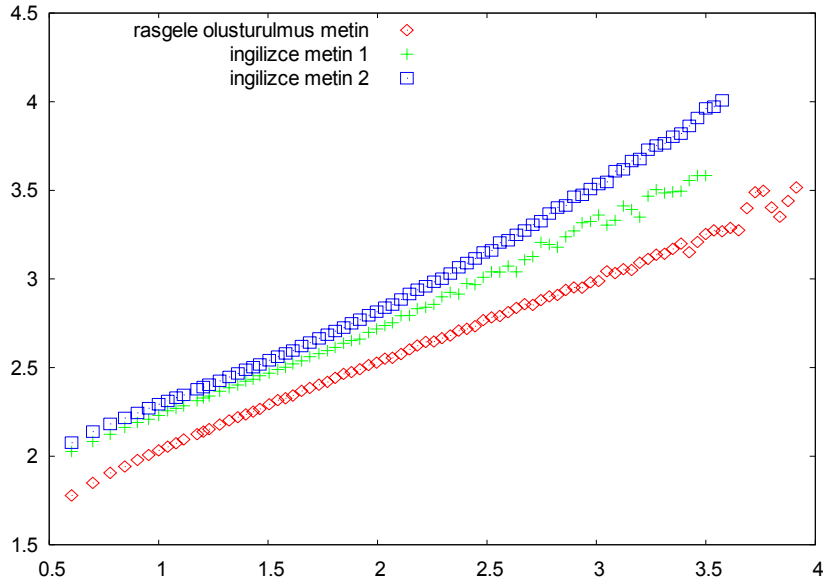
Yukarıda gözlenen korelasyonların metinlere değil de dillere ait olduğu savını güçlendirmek için aynı analiz önce Türkçe iki tamamıyla farklı metine uygulandı. Bu uygulama sonucunda aşağıda görülebileceği gibi iki metnin korelasyon yapısı birbirine benzer çıktı.



Aynı analizi içerik olarak birbirinden tamamıyla farklı İngilizce metinlere uyguladığımızda ise yine birbirine benzer korelasyon özellikleri görmekteyiz (bakınız aşağıda ki grafik).



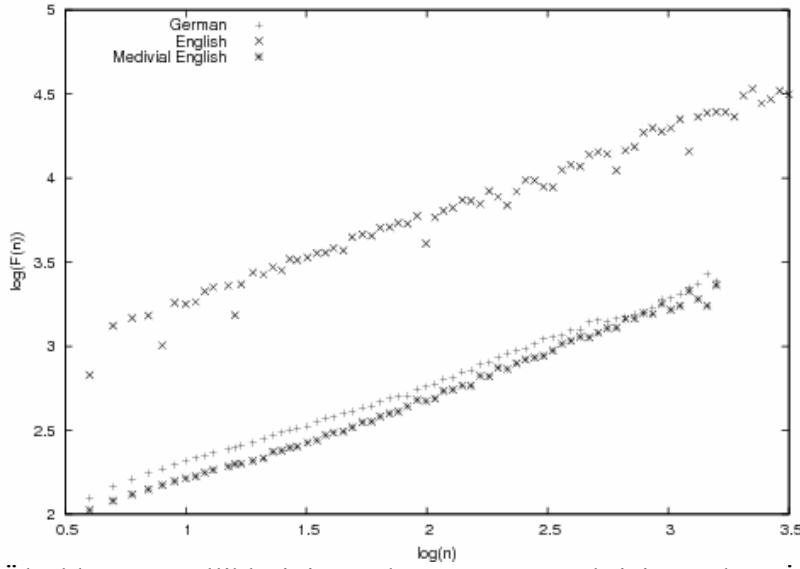
Gözlenen korelasyonların dillrin kendi yapısına özgü olduğu savını güçlendirmek için aynı analizi rasgele oluşturulmuş bir metne de uygulanması sonucunda rasgele oluşturulan metnin korelasyon özelliklerinin farklı olduğu sonucuna ulaşıldı.



Parametrizasyonun sağladığı sonuçların evrenselliğini görmek için aynı analiz başka dillere de uygulandı. Sonuçlar aşağıdaki tabloda sunulmuştur.

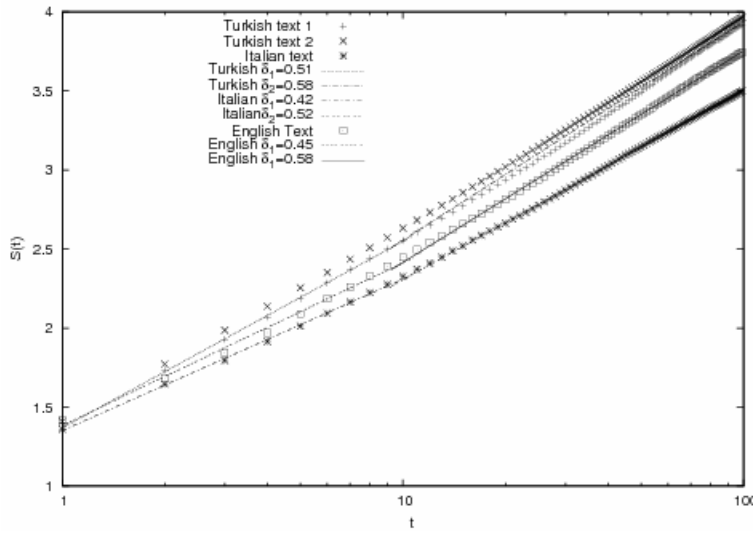
| Dil | 1. Bölge | 2.Bölge |
|------------|----------|---------|
| Arapça | 0.47 | 0.54 |
| Boşnakça | 0.48 | 0.47 |
| Fince | 0.52 | 0.77 |
| Türkçe | 0.52 | 0.67 |
| Portekizce | 0.48 | 0.7 |
| Fransızca | 0.49 | 0.64 |
| İtalyanca | 0.42 | 0.57 |
| İspanyolca | 0.49 | 0.56 |
| Lehçe | 0.47 | 0.57 |
| İndonesya | 0.50 | 0.63 |
| Malay | 0.51 | 0.66 |
| Zwahili | 0.49 | 0.41 |
| Almanca | 0.48 | 0.7 |
| Chaucer | 0.49 | 0.6 |
| İngilizce | 0.48 | 0.56 |

Aynı dil ailesine ait dillerin korelasyon özelliklerinin benzer olması oldukça dikkat çekicidir. Aslında İngilizcenin Almancaya daha çok benzemesi beklenirdi. Bunun tarihsel gelişimden kaynaklanabileceği düşünülerek araştırmaya orta çağ İngilizcesi de (Chaucer) dahil edildi. Ortaçağ İngilizcesinin beklendiği gibi Almancaya çok yakın çıktığı görüldü.

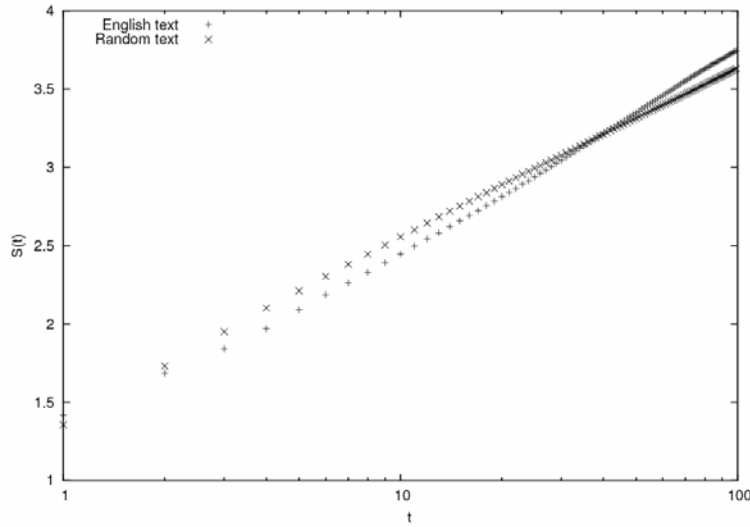


Ölçeklenme özelliklerinin sağlamasını yapmak için Türkçe, İtalyanca ve İngilizce metinler üzerinde Difüzyon Entropisi Analizi de kullanıldı. Difüzyon Entropisi Analizi sonucu elde edilen ölçeklenme üsteli δ ile gösterilmiştir. Burada da ölçeklenme katsayısının değişimi tüm diller için bir birinden farklı olarak gözlemlenebilmektedir. Türkçe metinler aynı δ değerleriyle tanımlanabilmekte ve her dil kendine has δ değerleri göstermektedir. (bkz. Aşağıdaki resim.) . Kaynak [21] ' e göre Difüzyon Entropisi Analizi ve Eğilimden Arındırılmış Dalgalanma Analizi kullanılarak zaman serilerinin Gausyen istatistiğe (normal ölçeklenme) göre mi yoksa Levy istatistiğine (anormal ölçeklenme) göre mi davrandığı belirlenebilir. Normal ölçeklenme durumunda Eğilimden Arındırılmış Dalgalanma Analizi'nin verdiği α üsseli Difüzyon Entropisi Analizinin verdiği δ üsseline eşit olmalıdır ($\delta = \alpha$). Diğer durumda ise $\delta = \frac{1}{3-2\alpha}$ olmalıdır. Aşağıda ki tabloda α_1 , α_2 , δ_1 ve anormal ölçeklenme durumunda α 'dan yukarıda ki bağıntıya göre hesaplanan δ değerleri bulunmaktadır. Bu tabloya göre Türkçe Levy rejiminde başlamış ve yine büyük ihtimalle orada kalmıştır ($\delta = \frac{1}{3-2\alpha}$). İtalyanca ise Gauss rejiminde başlayıp Levy rejimine geçmiş gibi görünmektedir.

| Dil | δ_1 | δ_2 | α_1 | α_2 | δ_1 | δ_2 |
|---------|------------|------------|------------|------------|------------|------------|
| Turkish | 0.51 | 0.58 | 0.52 | 0.67 | 0.51 | 0.60 |
| Italian | 0.42 | 0.52 | 0.42 | 0.56 | 0.46 | 0.53 |
| English | 0.45 | 0.58 | 0.48 | 0.56 | 0.49 | 0.53 |



Bunların yanı sıra difüzyon analizini rasgele oluşturulmuş İngilizce metnine de uyguladığımızda ortaya çıkan korelasyon özelliklerinin normal İngilizce metinlerden çok farklı olduğu ortaya çıkmıştır. Bu bir dile ait olan korelasyon özelliklerinin aynı zamanda anlamla bağlantılı olduğunu da göstermektedir.



5. SONUÇ

Bu çalışmada metinlerin analizi için dağarlardan bağımsız yeni bir parametrisasyon önerilmiştir. Bu parametrisasyon kullanılarak Türkçe, İngilizce ve 12 farklı dilin uzun zamanlı korelasyon özellikleri incelenmiştir. İncelenen tüm dillerin kendilerine has korelasyon özellikleri gösterdiği bulunmuştur. Aslında bu herhangi bir dilin karmaşık ölçeksiz bir ağda rassal yürüyüşle modellenebileceğini ima etmektedir. Bunun yanı sıra bir dilin korelasyon özelliklerinin o dilin anlam oluşturma yetisiyle yakından ilgili olduğu da gözlemlenmiştir (ki bu da yukarıda ki savı desteklemektedir).

Difüzyon analizi ve Eğilimden Arındırılmış Dalgalanma Analizinin işaret ettiği iki farklı ölçeklenme analizi dilleri ayırt etmek için kullanılabilir. Ayrıca her dil için bu özelliklerin gözlemlenmesi gramer olarak farklılıklarına rağmen tüm dillere ait evrensel bir özelliğe işaret etmektedir.

6. KAYNAKLAR

- [1] Montemurro MA, Pury PA. Long-range fractal correlations in literary corpora. 2002. Available from cond-mat/0201139.
- [2] Bhan, J., S. Kim, J. Kim, Y. Kwon, Long-range correlations in Korean literary corpora, *Chaos, Solitons and Fractals*, 29, 69–81, 2006.
- [3] H. E. Hurst, R. P. Black, Y. M. Simaika, *LongTerm Storage: An Experimental Study* (Constable, London)
- [4] Hrebicek, Ludek, *Text in Communication: Supra-Sentence structures*, Universitätsverlag Dr. N. Brockmeyer, Bochum, pp. 91-96, 1992.
- [5] Hrebicek, Ludek, *Quantitative Linguistics Vol. 566, Language Constructs, Constituents and the Menzerath-Altman Law*, Wissenschaftlicher Verlag Trier, 1995.
- [6] L. Hřebíček, G. Altmann (eds.): *Quantitative Text Analysis. (Quantitative Linguistics, Vol. 52)* Wissenschaftlicher Verlag Trier, Trier 1993, s.307
- [7] P. Allegrini, P. Grigolini and L. Palatella, Intermittency and scale-free networks: a dynamical model for human language complexity, *Chaos, Solitons and Fractals*, 20, 95-105, 2004.
- [8] Alexandre Rosas, Edvaldo Nogueira, Jr., and Jose F. Fontanari, Multifractal analysis of DNA walks and trails, *Physical Review E* 66, 061906, 2002.
- [9] C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Mosaic Organization of DNA Nucleotides, *Phys. Rev. E* 49, 1685, 1994.
- [10] N. Scafetta and P. Grigolini, Scaling detection in time series: Diffusion entropy analysis, *Phys. Rev. E*, 66, 036130, 2002.
- [11] S.V. Buldyrev, A.L. Goldberger, S. Havlin, R.N. Mantegna, M.E. Matsa, C.-K. Peng, M. Simons, H.E. Stanley, Long-Range Correlation Properties of Coding and Noncoding DNA Sequences: GenBank Analysis, *Phys. Rev. E* 51, 5084, 1995.
- [12] C.-K. Peng, S.V. Buldyrev, A.L. Goldberger, R.N. Mantegna, M. Simons, H.E. Stanley, Statistical Properties of DNA Sequences, *Physica A* 221, 180, 1995.

- [13] S.V. Buldyrev, N.V. Dokholyan, A.L. Goldberger, S. Havlin, C.-K. Peng, H.E. Stanley, G.M. Viswanathan, Analysis of DNA Sequences Using Methods of Statistical Physics, *Physica A* 249, 430, 1998.
- [14] C.-K. Peng, J. Mietus, J.M. Hausdorff, S. Havlin, H.E. Stanley, A.L. Goldberger, Long-range anticorrelations and non-Gaussian behavior of the heartbeat, *Phys. Rev. Lett.* 70, 1343, 1993.
- [15] C.-K. Peng, S. Havlin, H.E. Stanley, A.L. Goldberger, Quantification of Scaling Exponents and Crossover Phenomena in Nonstationary Heartbeat Time Series, *Chaos* 5, 82, 1995.
- [16] K.K.L. Ho, G.B. Moody, C.-K. Peng, J.E. Mietus, M.G. Larson, D. Levy, A.L. Goldberger, Predicting Survival in Heart Failure Case and Control Subjects by Use of Fully Automated Methods for Deriving Nonlinear and Conventional Indices of Heart Rate Dynamics, *Circulation* 96, 842, 1997.
- [17] G.M. Viswanatha, C.-K. Peng, H.E. Stanley, A.L. Goldberger, Deviations from uniform power law scaling in nonstationary time series, *Phys. Rev. E* 55, 845, 1997.
- [18] C.K. Peng, J.M. Hausdorff, S. Havlin, J.E. Mietus, H.E. Stanley, A.L. Goldberger, Multiple-time scales analysis of physiological time series under neural control, *Physica A* 249, 491, 1998.
- [19] Y.H. Liu, P. Cizeau, M. Meyer, C.-K. Peng, H.E. Stanley, Quantification of Correlations in Economic Time Series, *Physica A* 245, 437, 1997.
- [20] P. Cizeau, Y.H. Liu, M. Meyer, C.-K. Peng, H.E. Stanley, Volatility distribution in the S and P 500 stock index, *Physica A* 245, 441, 1997.
- [21] N. Scafetta, V. Latora and P. Grigolini, Levy scaling: The diffusion entropy analysis applied to DNA sequences, *Phys. Rev. E*, 66, 031906, 2002.
- [22] P. Allegrini, P. Grigolini, P. Hamilton, L. Palatella and G. Raffaelli, Memory beyond memory in heart beating, a sign of a healthy physiological condition, *Phys. Rev. E*, 65, 041926, 2002.
- [23] N. Scafetta, P. Hamilton, P. Grigolini, The thermodynamics of social process: the teen birth phenomenon, *Fractals* 9, 193 (2001).

